



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>5</sup> :</b> <b>C12N 15/00, 5/00, C12P 21/06</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 93/16177</b> <b>(43) International Publication Date:</b> 19 August 1993 (19.08.93)
<b>(21) International Application Number:</b> PCT/US93/00926 <b>(22) International Filing Date:</b> 2 February 1993 (02.02.93)  <b>(30) Priority data:</b> 07/833,992 11 February 1992 (11.02.92) US  <b>(71) Applicants:</b> CELL GENESYS, INC. [US/US]; 322 Lakeside Drive, Foster City, CA 94404 (US). THE PRESIDENT AND FELLOWS OF HARVARD COLLEGE [US/US]; 17 Quincy Street, Cambridge, MA 02138 (US).  <b>(72) Inventors:</b> SEIDMAN, Jonathan, G. ; 1350 Canton Avenue, Milton, MA 02186 (US). JAKOBOVITS, Aya ; 2021 Monterey Avenue, Menlo Park, CA 94025 (US).		<b>(74) Agents:</b> ROWLAND, Bertram, I. et al.; Flehr, Hobbach, Test, Albritton & Herbert, 4 Embarcadero Center, Suite 3400, San Francisco, CA 94111-4187 (US).  <b>(81) Designated States:</b> CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i>
<b>(54) Title:</b> HOMOGENOTIZATION OF GENE-TARGETING EVENTS  <b>(57) Abstract</b>  Homozygotic cells are obtained by employing homologous recombination with a construct comprising a marker gene. The marker gene allows for selection without amplification and by employing elevated levels of the antibiotic to which the marker gene imparts resistance, gene conversion can occur, where in a diploid host, both copies of the target locus will be the same. In this manner, knock-outs of genes can be readily achieved without requiring two steps of homologous recombination.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FR	France	MR	Mauritania
AU	Australia	GA	Gabon	MW	Malawi
BB	Barbados	GB	United Kingdom	NL	Netherlands
BE	Belgium	GN	Guinea	NO	Norway
BF	Burkina Faso	GR	Greece	NZ	New Zealand
BG	Bulgaria	HU	Hungary	PL	Poland
BJ	Benin	IE	Ireland	PT	Portugal
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SK	Slovak Republic
CI	Côte d'Ivoire	LI	Liechtenstein	SN	Senegal
CM	Cameroon	LK	Sri Lanka	SU	Soviet Union
CS	Czechoslovakia	LU	Luxembourg	TD	Chad
CZ	Czech Republic	MC	Monaco	TG	Togo
DE	Germany	MG	Madagascar	UA	Ukraine
DK	Denmark	ML	Mali	US	United States of America
ES	Spain	MN	Mongolia	VN	Viet Nam
FI	Finland				

## HOMOGENOTIZATION OF GENE-TARGETING EVENTS

### INTRODUCTION

#### Technical Field

The field of this invention is directed targeting of  
5 genes in viable cells.

#### Background

There are continuing advances in the ability to modify  
cells. One of the most exciting areas for the treatment of  
genetic diseases is gene therapy. In those situations where  
10 cells can be isolated, which are relatively long-lived and  
can be grown in culture and expanded, one has the  
opportunity to isolate such cells from a host, either the  
diseased host or an allogeneic host. The cells may then be  
grown in culture, expanded, and modified by homologous  
15 recombination, and the modified cells restored to the host.

In other situations, one wishes to produce products  
which are exogenous to the host. In this situation, one may  
modify embryonic stem cells, add them to a blastocyst, and  
then produce a chimeric host. This chimeric host may then  
20 be mated and the progeny either used as heterozygous mutants  
or further mated to produce homozygous mutants. By  
providing for appropriate changes in the embryonic cells,  
one can produce various human proteins in bovine milk and  
the like.

25 Alternatively, one may wish to change the nature or  
characteristics of a product of a host, such as meat,  
vegetable oils, size, or the like. In many cases,  
homologous recombination will be required with both genes,

-2-

since the host will be diploid. This means that one must go through an extensive number of steps in first modifying one gene, ensuring that the proper modification has occurred, then using the modified cells to repeat the process and  
5 modify the second gene. Even where the modification is a knock-out, namely introduction of an interfering sequence and/or deletion into the target locus, the process can be time-consuming and labor-intensive. There is, therefore, substantial interest in being able to develop techniques  
10 which allow for improvements in the efficiency of achieving genetic modifications at target loci.

#### Relevant Literature

Procedures for the modification of ES cells and detection of the modified cells may be found in Mortensen,  
15 et al., Proc. Natl. Acad. Sci. (USA) 88, 7036-7040 (1991). Paludan, et al., Gene 85, 421-426 (1989) reports enhanced resistance of cells comprising multiple neo genes. Spontaneous production of homozygous cells from heterozygous cultured cells have been reported by Nelson, et al., Mol.  
20 Cell. Biol. 9, 1284-1288 (1989); Potter, et al., Proc. Natl. Acad. Sci. (USA) 84, 1634-1637 (1987); Rajan, et al., J. Immunol. 145, 1598-1602 (1990) and Campbell, et al., Mol. Cell. Biol. 1, 336-346 (1981).

#### SUMMARY OF THE INVENTION

25 Homogenotization of gene targeting events is achieved by introducing DNA into cells, where the DNA has homology to a target locus and contains a marker gene. Cells undergoing homologous recombination are selected and subjected to elevated levels of the selection agent. Cells are then  
30 selected which have undergone gene conversion, where the mutant is now homozygous for the modification. The selective gene is normally one that imparts drug resistance and is normally incapable of amplification.

DESCRIPTION OF THE SPECIFIC EMBODIMENTS

Mutant cells which are homozygous at a targeted locus are produced by introducing DNA into the cells, where the DNA has homology to the target locus and includes a marker gene, allowing for selection of cells comprising the integrated construct. The cells are then screened for homologous recombinants. The homologous integrants are exposed to elevated levels of the marker gene, selecting for cells which have multiple copies of the selective agent by other than amplification. The cells are then analyzed for homozygosity at the target locus.

Various constructs can be prepared for homologous recombination at a target locus. Usually, the construct will include at least 50 bp of sequence homologous with the target locus and not more than 500 Kbp. The marker gene will normally be flanked on both sides by homologous sequences.

Various marker genes may be employed, which have the following characteristics. The marker gene is not amplifiable in that with increasing concentrations of the selective agent, one does not detect spontaneous events in which the gene has undergone tandem replication. The selective agent should not be extremely toxic, so that minor variations in concentration should not have significant effects on the viability of cells. Thus, one can provide for easy gradations in change of concentration of the selective agent so as to specifically select for cells having increased resistance to the selective agent. Gradations will be generally such that upon plating the heterozygous mutant at higher concentrations of the selective agent, there will colony forming efficiency of from about 1 in  $10^{-10}$  to  $10^{-5}$ .

The transcriptional initiation region for the marker gene may be any convenient region, either inducible or constitutive. By providing for an inducible transcription initiation region, the marker gene may be silent when the cells are introduced into a host, where the inducible agent is absent or the gene is turned on upon being introduced

into a host or upon maturation and differentiation of the cells. Inducibility may be as a result of temperature, specific compounds, etc. A wide variety of promoters have been described in the literature, which are constitutive or  
5 inducible, where induction may be the result of temperature, inducing agents, any specific cell type or specific level of maturation. Also, a number of viral promoters are known, which may also find use. Promoters of interest include the  $\beta$ -actin promoter, SV40 early and late promoters,  
10 phosphoglycerate kinase promoter ("PGK"), immunoglobulin gene promoter, thymidine kinase promoter, human cytomegalovirus promoter and the Friend spleen focus-forming virus promoter. The promoters may or may not be associated with enhancers, where the enhancers may be naturally-  
15 ~~associated with the particular promoter or associated with a different promoter.~~

A termination region will be present, where the termination region may be naturally-associated with the marker gene or may be derived from a different gene. For  
20 the most part, the termination regions are not critical and a wide variety of termination regions may be employed without adversely affecting expression or the utility of the marker gene.

DNA constructs will be prepared for integration into  
25 the target locus. The construct will comprise the marker gene, with its appropriate transcriptional and translational initiation and termination regulatory regions, flanked by regions of homology with the target locus. The integration of the subject marker gene at the target locus may interfere  
30 with functional expression in a variety of ways, preferably by insertion into the coding region or an exon of the target locus, but may also interfere by insertion into an intron where the construct has a splice-acceptor site adjacent to the marker gene, so as to truncate the target locus and  
35 prevent expression of a functional protein. If one wishes, rather than providing for an independent transcriptional initiation region as part of the construct, one could provide for integration into the target locus at a site

which introduces the marker gene under the transcriptional control of the transcription initiation region of the target gene, e.g. into an intron where the construct has at the 5' terminus of the marker gene a splice-acceptor site, or into an exon resulting in fusion of the reading frames of the target and marker genes. The significant factor is that integration of the marker gene at the target locus results in inactivation of the target gene.

The regions of homology with the target locus will usually be at least about 50 bp, more usually at least about 150 bp and may be as high as 500 Kbp, usually not more than about 20 Kbp, preferably not more than about 10 Kbp. Usually, there will be at least about 20, more usually at least about 50 bp of homology 5' and 3' of the marker gene. Various considerations will be involved in the extent of homology, such as the size of the target locus, availability of sequences, relative efficiency of double cross-over events at the target locus, similarity of the target sequence with other sequences, and the like.

The region(s) of homology may include mutations, where mutations may further inactivate the target gene, in providing for a frame shift, or changing a key amino acid, or the mutation may correct a dysfunctional allele, etc. Usually, the mutation will be a subtle change, not exceeding about 5% of the homologous flanking sequences, usually fewer than 50 bp. Where mutation of a gene is desired, the marker gene may be inserted into an intron, so as to be excised from the target gene upon transcription.

The construct will be prepared in accordance with known ways, where various fragments will be brought together, introduced into appropriate vectors, cloned, analyzed and then manipulated further until the desired construct has been made. Various modifications may be made to the sequence, to allow for restriction analysis, excision, identification of probes, or the like. Silent mutations may be introduced, as desired. At various stages, restriction analysis, sequencing, amplification with the polymerase

chain reaction, primer repair, *in vitro* mutagenesis, or the like may be employed.

The construct will usually be prepared using a bacterial vector, including a procaryotic replication system, e.g. an origin recognizable by E. coli, where at each stage, the construct may be cloned and analyzed. A marker, the same as or different from the marker to be used for insertion, may be employed, where the procaryotic marker is usually removed prior to introduction into the target cell. Once the vector containing the construct has been completed, it may be further manipulated by deletion of the bacterial sequences, linearization, introducing a short deletion in the homologous sequence, generally not exceeding about 500 bp, generally below about 20% of the homologous sequence, usually in the range of about 50-300 bp. After final manipulation, the construct is ready to be introduced into the cell.

A wide variety of eukaryotic target host cells may be employed, normally cells from vertebrates, more particularly, mammals, desirably domestic animals or primates, particularly humans. The cells which may be subjected to transformation may be any mammalian cells of interest, which may find use in cell therapy, research, interaction with other cells *in vitro* or the like. Cells of particular interest include, among other lineages, stem cells, e.g. hematopoietic stem cells, embryonic stem cells, etc., the islets of Langerhans, adrenal medulla cells which may secrete dopamine, osteoblasts, osteoclasts, epithelial cells, endothelial cells, leukocytes, e.g. B- and T-lymphocytes, myelomonocytic cells, etc., neurons, glial cells, ganglion cells, retinal cells, liver cells, e.g. hepatocytes, bone marrow cells, keratinocytes, hair follicle cells, and myoblast (muscle) cells.

For embryonic stem cells, an embryonic stem cell line may be employed or embryonic stem cells may be obtained freshly from a host. The cells may be grown on an appropriate fibroblast fetal layer or grown in the presence of leukemia inhibiting factor (LIF) and then used. The



embryonic stem cells may be injected into a blastocyst to provide a chimeric animal.

Upstream and/or downstream from the target gene construct may be a gene which provides for identification of whether a double cross-over has occurred. For this purpose, the herpes simplex virus thymidine kinase ("HSV-tk") gene may be employed, since the presence of the thymidine kinase gene may be detected by the use of nucleoside analogs, such as acyclovir or gancyclovir, for their cytotoxic effects on cells that contain a functional HSV-tk gene. The resistance to these nucleoside analogs indicates the absence of the thymidine kinase gene and, is therefore, indicative that a double cross-over event has occurred and thus homologous recombination has also occurred.

Various techniques may be employed for introduction of the construct into the host cell. Techniques which may be used include calcium phosphate/DNA coprecipitates, micro-injection of DNA into the cell, electroporation, lipofection, bacterial protoplast fusion with intact cells, transfection or the like. The DNA may be single- or double-stranded, linear or circular, relaxed or supercoiled DNA. For various techniques for transforming mammalian cells, see Keown, et al., Meth. in Immun. 185, 527-537 (1990). The cells may then be grown in appropriately-selected medium to identify cells providing the appropriate integration. Those cells which show the desired phenotype may then be further analyzed by restriction analysis, electrophoresis, Southern analysis, polymerase chain reaction, or the like. By identifying fragments which show the appropriate insertion at the target gene site, one can identify cells in which homologous recombination has occurred to inactivate or otherwise modify one of the two copies of the target gene.

The heterozygote cells containing the integrated marker gene may then be expanded in an appropriate culture medium over an extended period of time, usually at least one week and up to six weeks or more. The cells may then be plated, and individual colonies selected at varying concentrations of the drug to which the marker gene imparts resistance.

For example, with neomycin resistance, depending upon the nature of the cell, the concentration of G418 may vary from about 0.5-4 mg/ml, more usually from about 0.5-2.5 mg/ml. The surviving colonies may then be screened to establish  
5 that the functional endogenous gene is absent or has been properly mutated. As described above, various techniques may be used for analyzing for the presence of homozygosity of the mutant. The resulting homozygote will have two copies of the same gene on different chromosomes, normally  
10 including the marker gene. The cell usually will otherwise be free of the construct.

The subject invention may find wide application, where one wishes to inactivate a particular gene or introduce an alteration into a gene at a target locus. Genes of interest  
15 include a variety of ~~surface membrane protein genes~~. One area of interest is reduction in the expression of major histocompatibility complex antigens, Class I and/or II, or any of the individual groups, such as A, B, C, DP, DQ and DR. For knocking out Class I, one may provide for  
20 integration into the  $\beta$ 2-microglobulin gene. For Class II antigens, it may be sufficient to knock out the DR genes. Thus, one may reduce the immunologic response associated with the presence of foreign MHC antigens, when such antigens are introduced into a host. Alternatively, one may  
25 wish to inhibit expression of the host gene when one is introducing an exogenous immunoglobulin gene, e.g. knocking out the host immunoglobulin locus, when introducing an exogenous locus. Other genes of interest include those affecting MHC expression or other components of the antigen  
30 presentation machinery of the cell which ultimately affect immunogenicity.

One may also use the subject method to study various metabolic processes, to identify the effect of the absence of a particular genetic capability, and the like. In this  
35 manner, one may be able to identify the role of particular oncogenes, the effect of regulatory proteins, the roles of enzymes, the presence of alternative pathways, and the like.

One may also use the subject methodology for introducing mutations into dysfunctional alleles, so as to cure genetic defects, particularly where subtle changes are involved such as those involving fewer than about 50 bp, frequently fewer than about 10 bp. Thus, one could modify stem cells, such as hematopoietic stem cells, to cure diseases such as sickle cell anemia,  $\beta$ -thalassemia, hemophilia, etc., targeting genes for gene therapy which include genes for  $\beta$ -globin, enzymes of erythrocyte metabolism, the complement system, coagulation factors, dystrophin, enzymes of carbohydrate, lipid, amino acid, steroid and purine and pyrimidine metabolism, transport proteins, e.g., cystic fibrosis transmembrane regulator, and the like.

The modified embryonic stem cells may be used to produce chimeric non-human mammals by injecting the homogenotized embryonic stem cells into a blastocyst and then brought to term in a female host mammal in accordance with conventional techniques. Heterozygous progeny may then be screened for the presence of the alteration at the site of the target locus, using PCR, Southern blotting or the like. After mating with a wild-type host of the same species, the resulting chimeric progeny may then be cross-mated to achieve homozygous hosts. The homozygous hosts may have a variety of alterations, where the gene at the target locus may be made non-functional, may be retained as functional, but altered, or may be retained and additional genetic capability provided, where the additional capability may be coregulated by concomitant transcription with the gene at the target locus.

The mammalian host will have target loci comprising alleles which should be at least substantially identical, in having the marker gene integrated at the same position at both alleles. In addition, any alterations as a result of the homologous recombination in the first step should also be present in both alleles. In this way a "double knockout," symmetrical integration, which may include one or

more alterations at the target locus, or the like may be achieved.

The homozygotes may be subject to further genetic modification. For example, one may wish to introduce  
5 additional genetic capability into the homozygotic hosts, where the endogenous alleles have been made non-functional, to substitute, replace or provide different genetic capability to the host. For example, one may wish to produce human immunoglobulins in small host, such as a mouse  
10 or rabbit. By first knocking out the endogenous genes encoding for the light and heavy chains of the immunoglobulins, one can then introduce genes from the human loci, which genes encode for the human light and heavy chains and are capable of rearrangement in response to  
15 immunization of the host to produce human immunoglobulin proteins. For example, one may knockout the J regions, which prevents expression of the immunoglobulin subunits.

One may wish to remove the marker gene after homogenotization. By introducing a construct comprising  
20 substantially the same homologous DNA, possibly extended sequences, having the marker gene portion of the original construct deleted, one may be able to obtain homologous recombination with the target locus. By using a combination of marker genes for integration, one providing positive  
25 selection and the other negative selection, in the removal step, one would select against the cells retaining the marker genes. Convenient marker genes for negative selection include the herpes simplex virus thymidine kinase gene, whose presence results in sensitivity to nucleoside  
30 analogs, such as acyclovir and gancyclovir.

In bovine or ovine hosts, one may wish to replace the bovine or ovine protein with a human protein. In domestic animals, one may wish to introduce alterations which result in improved products or growth of the host. Where domestic  
35 animals are prone to diseases as a result of particular alleles, one may alter the alleles to reduce the susceptibility to the disease.

-11-

Hosts of interest include, primates other than man, ovine, bovine, equine, canine, feline, lagomorpha, murine, etc.

The following examples are offered by way of illustration and not by way of limitation.

### EXPERIMENTAL

#### EXAMPLE I

Four separate genes, G-protein subunits  $\alpha_2$  and  $\alpha_3$ , T-cell receptor (TCR)- $\alpha$ ,  $\beta$ -cardiac myosin heavy chain (MYHC-B) were targeted using positive/negative selection (Mansour, et al., Nature 336, 348-352 [1988]). In each case, a targeting construct was made containing the neomycin-resistance gene (neo) driven by the promoter region of phosphoglycerate kinase (PGK) gene. The PGK-neo sequence interrupted an exon creating a detectable change in the gene's restriction enzyme digestion pattern.

In each case PGK-TK phosphoglycerate kinase promoter-thymidine kinase gene) was inserted outside the regions of homology. A) The production of the  $\alpha_2$  target construct has been described previously in Mortensen, et al., Proc. Natl. Acad. Sci. (USA) 88, 7036-7040 (1991). B)  $\alpha_3$  gene clone was isolated as described for the  $\alpha_2$  gene. A Hind III-Sal I fragment was subcloned into Bluescript SK (+) (Stratagene). The construct contained the PGK-neo gene interrupting exon 1 at an Nco I site at the translational start ATG. C) A 5.6 Kbp fragment of the murine TCR- $\alpha$  gene containing all 4 constant-region exons was subcloned from genomic clone  $\lambda$ \*1.2 into Bluescript SK II (+) (Ausubel, et al., Curr. Proto. in Mol. Biol. [John Wiley and Sons, New York, 1988]). The fragment extends from an artificial EcoRI site, 5' of the Xba I site to the XhoI site. The EcoRI site between exons 3 and 4 was deleted by digestion with EcoRI, blunting with Klenow and subsequent ligation. The first constant-region exon was interrupted at a unique EcoRV site by the insertion of PGK-neo. D) A 7.5 Kbp (SalI-EcoRV) portion of the Myhc-B gene extending from approximately exon 10 to approximately exon 24 was cloned from a Balb/c genomic library (Clontech),

identity was confirmed by sequencing. A targeting construct was made by interrupting the gene at a unique Nde I site in exon 14.

ES cells (line CCE obtained from E. Robertson, Columbia University or CC1.2 obtained from A. Bradley, Baylor University) were cultured, transfected and screened for homologous recombination as previously described (Mortensen, *supra* [1991]). Several cell lines identified as heterozygous for homologous recombination were expanded for 14-28 days. Cell lines were then plated, cultured for 24 hours and then selected at varying concentrations of G418. DNA from the surviving clones was analyzed by Southern blot for absence of the endogenous gene.

Table 1. Loss of heterozygosity in ES cell lines

Gene (Parent Line)	Cell Line	No. Cells Plated*	[G418] (mg/ml)			
			1.0-1.5		2.0	
			No. of Colonies	% Homozygote (No. Analyzed)	No. of Colonies	% Homozygote (No. Analyzed)
$\alpha_{12}$ (CCE)	17E10	$2 \times 10^4$	23	100% (2)	2	0% (2)
	18D3	$2 \times 10^4$	175	12% (26)	40	ND <sup>#</sup>
	32-32	$1 \times 10^5$	100	53% (17)	17	88% (17)
	32-28	$1 \times 10^6$	27	4% (28)	0	
$\alpha_{13}$ (CC1.2)	1A4	$5 \times 10^5$	37	43% (23)	9	44% (9)
TCR- $\alpha$ (CCE)	6-6	$5 \times 10^5$			54	24% (54)
	6-22	$5 \times 10^5$	76	7% (76)	2	50% (2)

\* Efficiency of plating 40-80%  
# ND - not determined

Southern blot analysis of DNA from the clones isolated in high concentrations of G418 was carried out. DNA was fractionated on 1% agarose gels, transferred to Genescreen and hybridized as previously described (Ausubel, *supra* [1988]). In each case, a band was observed for the homologous recombination product, for  $\alpha_{12}$ , a 13.6 Kb band, for  $\alpha_{13}$ , a 6.5 band, for TCR- $\alpha$ , a 2.0 band and for Myhc-B, a 7.1 band. The  $\beta$ -cardiac myosin probe cross-reacted with the

-13-

Myhc-A ( $\alpha$ -cardiac myosin heavy chain), gene resulting in 2 bands being present in each lane.

To determine the frequency of homozygous mutant ES cells surviving high G418, 4 individual heterozygous subclones were isolated by plating trypsin-dispersed 17E10 cells at high dilution. Each of these subclones and the original 17E10 clones were plated  $5 \times 10^5$  cells/100 mm plate, selected at high G418 levels and the surviving clones analyzed as in Table 1. Plating efficiency without G418 was 42%.

Table 2. Frequency of homozygous mutant ES cells surviving high G418

[G418] (mg/ml)					
Subclone of 17E10	1.0		2.0		Apparent No. of Colonies#
	No. Colonies	% Homozygote (No. Analyzed)	No. Colonies	% Homozygote (No. Analyzed)	
1	143	75% (20)	28	94% (18)	114.4
2	152	92% (24)	52	95% (19)	139.3
3	44	22% (9)	0		9.8
4	40	95% (21)	5	100% (4)	38.1
Average ± Std. Dev.					75.4± ± 61.4

\* Calculated from (fraction of homozygotes at 1 mg/ml)  $\times$  (total no. of colonies)

† An estimated rate of mutation per cell generation of  $1.3 \times 10^{-5}$  was calculated from the number of generations since recloning (assuming a generation time of 14.2 hours<sup>2</sup> and the fraction of homozygotes ((average No. of colonies)/((No. of cells plated)  $\times$  (plating efficiency))) according to the equation rate = (fraction of homozygotes)/(No. of generations since cloning). Cells were cultured for 16 days (27 generations) after plating at high dilution. Mass cultures from the original clone were selected in parallel giving 241, 261, 228, and 262 colonies per 100 mm plate containing >90% homozygotes (18 analyzed).

## EXAMPLE II

### Inactivation of the mouse Ig heavy chain J genes in ES cells.

#### Construction of the inactivation vector

A 6.1-Kb EcoRI fragment, containing the mouse immunoglobulin heavy chain J region genes and flanking sequences, cloned from a Balb/c mouse embryo genomic library and inserted into pUC18 (pJH), was digested with XhoI and NaeI to delete an about 2.3 kbp fragment containing the four J genes (see Chart 2A). An about 1.1 kbp XhoI-BamHI fragment, blunted at the BamHI site, containing a neomycin

-14-

resistance gene driven by the Herpes simplex virus thymidine kinase gene (HSV-tk) promoter and polyoma enhancer was isolated from pMC1Neo (Thomas and Capecchi, Cell, 51, 503-512, 1987). This fragment was inserted into the

5 XhoI-NaeI deleted pJH to form the inactivation vector (pmH6J, see Chart 2B), in which the transcriptional orientation of the neomycin and the heavy chain genes is the same. This plasmid was linearized by NdeI digestion before

10 homologous recombination event are about 2.8 kbp and about 1.1 kbp fragments, located 5' and 3' to the neomycin gene, respectively.

#### Culturing, Electroporation, and Selection of ES cells

The ES cell line E14TG2a (Koller and Smithies, 1989,

15 PNAS, USA, 86, 8932-8935) was cultured on mitomycin C-treated embryonic fibroblast feeder layers as described (Koller and Smithies, 1989, PNAS USA, 86, 8932-8935). ES cells were trypsinized, resuspended in HBS buffer (pH 7.05; 137mM NaCl, 5mM KCl, 2mM CaCl<sub>2</sub>, 0.7mM Na<sub>2</sub>HPO<sub>4</sub>, 21mM HEPES

20 pH 7.1) at a concentration of  $2 \times 10^7$ /ml and electroporated in the presence of 50  $\mu$ g/ml of the linearized inactivation vector. Electroporation was carried out with a BioRad Gene Pulser using 240 volts and 500 $\mu$ F capacitance.  $5 \times 10^6$  electroporated cells were plated onto mitomycin C-treated

25 fibroblasts in 100mm dishes in the presence of Dulbecco's modified Eagle's media (DMEM) supplemented with 15% fetal bovine serum and 0.1 mM 2-mercaptoethanol. The media was replaced 24 hr after electroporation with media containing 200 $\mu$ g/ml G418. G418-resistant ES colonies resulting 12-14

30 days after electroporation were picked with drawn out capillary pipettes for analysis using the polymerase chain reaction (PCR). Half of each picked colony was transferred to an individual well of a 24-well plate, already seeded with mitomycin C-treated feeder cells. The other halves,

35 combined in pools of four, were transferred to Eppendorf tubes containing 0.3 ml of PBS and cell lysates were prepared for PCR analysis as described by Joyner et al



(Nature, 338:153-155, 1989). The PCR reaction included 5-20 $\mu$ l of the cell lysate, 1 $\mu$ M of each primer, 1.5u of Taq polymerase and 200 $\mu$ M of dNTPs. The PCR amplification employed 45 cycles using a thermal cycler (Perkins-Elmer Cetus), with 1 min. melt at 94°C, 2 min. annealing at 55°C, and 3 min. extension at 72°C. The two priming oligonucleotides are ACGGTATCGCCGCTCCCGAT (SEQ ID NO:1) and AGTCACTGTAAAGACTTCGGGTA (SEQ ID NO:2), which correspond respectively to about 120 bases 5' of the BamHI site of the neomycin gene, and to the sequences located in the mouse heavy chain gene, about 160 bases 3' of the insertion site. Successful homologous recombination gives rise to an about 1.4 kbp fragment. 20 $\mu$ l of the reaction mixture is electrophoresed on 1% agarose gels, stained with ethidium bromide and transferred to nylon membranes (Gene Screen). Filters were probed with a <sup>32</sup>P-labelled EcoRI-PstI about 1.4 kbp fragment located in the mouse heavy chain, 3' of the insertion site (see Chart 2). For further analysis, genomic DNA was prepared from ES cells, digested with restriction enzymes as recommended by the manufacturers, and fragments were separated on 1% agarose gels. DNA was transferred to nylon membranes (Gene Screen) and probed with the <sup>32</sup>P-labelled fragment as described above.

#### Analysis of G418-resistant ES colonies

In the first experiment, PCR analysis of the pooled colonies detected one positive PCR signal of the expected size (about 1.4kbp) out of 34 pools representing 136 G418-resistant colonies. The four individual colonies that had contributed to this positive pool were analyzed individually by PCR, and a positive clone, ES33D5, was identified. Similar analysis of 540 G418-resistant colonies obtained in the second experiment yielded 4 additional positive clones (ES41-1), ES61-1, ES65-1, ES110-1).

In order to verify the targeting disruption of one copy of the J genes, (the gene is autosomal and thus present in two copies), the PCR positive clones were expanded and genomic DNA was prepared, digested with HindIII or with SacI

and analyzed by Southern analysis as described using the EcoRI-PstI probe.

The replacement of the J genes by insertion of the neomycin gene by an homologous recombination event results in an HindIII fragment, detectable with the EcoRI-PstI probe, which is about 1.9 kbp longer than the equivalent fragment in the native locus, due to the loss of two HindIII sites located in the deleted J gene region (see Chart 2C). Southern analysis of each of the 5 positive clones by HindIII digestion gave a pattern which indicated that one of the two copies of the heavy chain J genes had been disrupted. Three labeled fragments were detected: one fragment (about 760 bp), identical in size to that present in untreated cells at the same intensity, one fragment (about 2.3kbp) identical in size to that present in untreated cells, but of decreased intensity in the PCR positive clone, and an additional fragment about 4.2 kbp, the size predicted for an homologous recombination event, present only in the PCR-positive clones. Similarly, the replacement of the J genes by the neomycin gene by an homologous recombination event results in a loss of one SacI site and the appearance of a fragment, detectable with the EcoRI-PstI probe, which is about 570 bp smaller than the equivalent fragment in the native locus (see Chart 2C). Southern analysis of the clones by SacI digestion gave the expected pattern of one native and one targeted allele: about 4.0 kbp fragment, identical in size to that detected in untreated cells, but of decreased intensity in the 5 positive clones, and an additional fragment of about 3.4 kbp, the size predicted for a targeted homologous recombination event, present only in the identified clones. Rehybridization of the Southern blots with a probe for the neomycin gene shows that only the 4.2 kbp and 3.4 kbp fragments, resulting from the HindIII and the SacI digestion, respectively, hybridized to the probe as predicted by the targeting event.

-17-

Inactivation of mouse immunoglobulin heavy chain J  
genes in mice

Injection of targeted ES cells into mouse blastocysts  
and generation of chimeric offsprings

5 Mice were purchased from Jackson Laboratories (Bar Harbor, ME). Three and a half day old C57BL/6 blastocysts were obtained from 4-5 week old superovulated females as described by Koller et al. 1989 *supra*. ES cells were trypsinized, washed once with fresh DMEM media and diluted  
10 to about  $1 \times 10^6$ /ml in M2 media. About  $5 \mu\text{l}$  of cells were added to a  $150 \mu\text{l}$  droplet of M2 media, under paraffin oil, containing the blastocysts. Ten to fifteen cells were injected into the blastocoel of each blastocyst. Six to  
15 nine ES cell-containing blastocysts were returned to each uterine horn of C57BL/6 x DBA F1 pseudopregnant females mated 2.5 days previously with vasectomized males. Pups derived from the injected blastocysts were generally born 16-18 days later. The contribution of the ES cells to the  
20 offspring was judged visually by examination of the coat color of the pups. The blastocysts were obtained from C57BL/6 mice, which are solid black in color. The ES cell line E14TG2a, the parental line from which the targeted cell lines were derived, was isolated from 129/Ola mice. This mouse strain is cream in color, the combined effect of three  
25 color genes, the dominant  $A^w$  allele at the agouti locus, the recessive pink-eyed-diluted allele at the p locus and the recessive  $C^{ch}$  allele at the C locus. Offspring in which the ES cells participated in the formation of the animal had coats containing brown and cream hair. The ES cell line  
30 ES41-1 carrying inactivated mouse immunoglobulin heavy chain, was injected into C57BL/6 mouse blastocysts as described above. Six out of the 18 surviving pups had a high degree of coat color chimerism (70-90%). PCR analysis of DNA isolated from chimeric newborn pups from a female  
35 implanted with blastocysts injected with the inactivated ES cells, indicated that the mutated immunoglobulin heavy chain locus is present in a variety of organs such as spleen, thymus, kidney, liver, brain and skin.

Homogenotization of inactivated Ig heavy-chain J genes in ES cells

Heavy-chain J gene-targeted ES 110-1 cell line was plated on feeder cells at the density of  $5 \times 10^3$  cells per 100mm plate in the presence of 1400  $\mu\text{g/ml}$  of G418. Seven of the surviving colonies were expanded, genomic DNA was prepared and analyzed by Southern analysis for the absence of endogenous Ig heavy-chain J gene. Southern analysis by Sac I digestion indicated that one of the clones, ES DK207, had lost the native Ig heavy chain locus. Hybridization of the same blot with an Xho I-Nae I probe, which spans the whole J gene region, did not detect any fragments in the ES DK207 genome, further indicating the loss of the native Ig heavy-chain allele.

15 Generation of chimeric mice carrying homogenotized heavy-chain targeted B cells

ES DK207 cells were injected into mouse blastocysts and chimeric mice were generated as described above. Then chimeric mice were obtained, seven of which had 40-70% coat color chimerism, indicating the ES cell contribution to these mice.

The mice were used to determine the effect of the deletion of the J genes from both heavy-chain alleles on the capability of the B cells to differentiate and express antibodies. Analysis of peripheral blood lymphocytes derived from one of these chimeras, DK334-1, indicated that about 10% of the lymphocytes are derived from the mouse strain from which the ES cells originated, namely the 129Sv/Ola strain. This was determined with Ly9a antibodies which bind to cells from strain 129, but not cells from strain C57B/6. The absence of mature B cells originating from strain 129 was established by demonstrating by means of fluorescence activated cell sorting that the lymphocytes from strain 129 lacked the surface antigens CD45R<sup>bright</sup> and IgM<sup>a</sup>. These results support the conclusion that due to the deletion of the J region in both alleles of the Ig heavy

chain, B cell development is inhibited in lymphocytes derived from ES DK207.

It is evident from the above results, that the subject methodology provides for an effective efficient way to  
5 produce homozygosity in two or more alleles in a host, where one wishes to change the nature of one or more alleles. In this manner, one may knock out the two alleles present in a diploid host, mutate two alleles, so as to have the same sequence, while still retaining a sufficiently high  
10 frequency to allow for isolation of the desired homozygotes. The subject method can be used in a wide variety of applications, in transgenic animals, modification of the surface proteins present on a host cell, so as to reduce or enhance immunogenicity, and the like.

15 All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

20 Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications may  
25 be made thereto without departing from the spirit or scope of the appended claims.

-20-

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Cell Genesys, Inc.  
The President and Fellows of Harvard College
- 5 (ii) TITLE OF INVENTION: Homogenization of gene-targeting events
- (iii) NUMBER OF SEQUENCES: 2
- (iv) CORRESPONDENCE ADDRESS:
- 10 (A) ADDRESSEE: Bertram I. Rowland  
(B) STREET: 4 Embarcadero Center, Suite 3400  
(C) CITY: San Francisco  
(D) STATE: California  
(E) COUNTRY: USA  
(F) ZIP: 94111
- 15 (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- 
- 20 (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:  
(B) FILING DATE: 02-FEB-1993  
(C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- 25 (A) NAME: Rowland, Bertram I.  
(B) REGISTRATION NUMBER: 20,015  
(C) REFERENCE/DOCKET NUMBER: FP-55841/BIR
- (ix) TELECOMMUNICATION INFORMATION:
- 30 (A) TELEPHONE: (415) 781-1989  
(B) TELEFAX: (415) 398-3249

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
- 35 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

40 ACGGTATCGC CGCTCCCGAT  
20

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
- 45 (A) LENGTH: 23 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA

-21-

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGTCACTGTA AAGACTTCGG GTA

23

WHAT IS CLAIMED IS:

1. A method for modifying eukaryotic cells at a multicopy target locus, said method comprising:
  - 5 introducing into host cells under integrating conditions a construct, said construct comprising a marker gene flanked by regions of homology with said target locus, to produce a mixture of cells comprising modified cells in which said construct is homologously recombined at a single target site, wherein said marker gene is a non-amplifiable  
10 gene;
  - selecting cells comprising said construct by means of ~~said marker gene by growing said cells in a selective medium~~ at a first level of selective agent and analyzing viable cells for homologous recombination to provide a  
15 substantially pure population of modified cells;
  - subjecting said population of modified cells to a level of selective agent greater than said first level, whereby gene conversion occurs to provide cells homozygotic at said target locus; and  
20 isolating said homozygotic cells.
2. A method according to Claim 1, wherein said marker gene is an antibiotic resistance gene.
3. A method according to Claim 2, wherein said antibiotic resistance gene is a neomycin resistance gene.
- 25 4. A method according to Claim 1, wherein said marker gene is flanked by at least 50bp of sequence homologous with said target locus.



-23-

5. A method according to Claim 4, wherein said homologous sequence is mutated from said target locus sequence.

6. A method according to Claim 1, wherein said host  
5 cells are embryonic stem cells.

7. A method for modifying mammalian cells at a multicopy target locus, said method comprising:

introducing into mammalian host cells under integrating  
conditions a construct, said construct comprising a marker  
10 gene flanked by regions of homology with said target locus,  
to produce a mixture of cells comprising modified cells in  
which said construct is homologously recombined at a single  
target site, wherein said marker gene is an antibiotic  
resistance gene;

15 selecting cells comprising said construct by means of  
said marker gene by growing said cells in a selective medium  
at a first level of antibiotic and analyzing viable cells  
for homologous recombination to provide a substantially pure  
population of modified cells;

20 expanding said pure population of modified cells;

subjecting said expanded population of modified cells  
to a level of antibiotic greater than said first level,  
whereby gene conversion occurs to provide cells homozygotic  
at said target locus; and

25 isolating said homozygotic cells.

8. A method according to Claim 7, wherein said  
antibiotic resistance gene is a neomycin resistance gene.

-24-

9. A method according to Claim 7, wherein said marker gene is flanked by at least 50bp of sequence homologous with said target locus.

10. A method according to Claim 9, wherein said homologous sequence is mutated from said target locus sequence.

11. A method according to Claim 7, wherein said host cells are embryonic stem cells.

12. A method according to Claim 7, wherein said target locus is a surface membrane protein locus.

13. A mammalian cell comprising two copies of the same endogenous gene into which is integrated at the same site a marker gene.

14. A mammalian cell according to Claim 13, wherein said marker gene is an antibiotic resistance gene.

15. A mammalian cell according to Claim 14, wherein said antibiotic is neomycin or G418.

16. A mammalian cell according to Claim 13, wherein said mammalian cell is an embryonic stem cell.

17. A chimeric mammal other than human comprising: homozygotic alleles comprising a marker gene encoding a product imparting antibiotic resistance to said chimeric mammal.

18. A chimeric mammal according to Claim 17, wherein said allele expresses a functional product.

19. A chimeric mammal according to Claim 17, wherein said allele is non-functional in being incapable of

-25-

expressing a product or expressing a non-functional product different from the wild-type product.

20. A chimeric mammal according to Claim 17, wherein said marker gene encodes for a product imparting G418  
5 resistance.

21. A chimeric mammal other than human comprising:  
non-functional homozygotic immunoglobulin J loci, being  
non-functional as a result of integration of a marker gene  
encoding a product imparting antibiotic resistance to said  
10 chimeric mammal.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US93/00926

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(5) : C12N 15/00, 5/00; C12P 21/06  
US CL : 800/2; 435/ 69.1, 172.3, 240.2, 320.1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
U.S. : 800/2; 435/ 69.1, 172.3, 240.2, 320.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)  
APS, Chemical Abstracts

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Nature, Volume 342, issued 23 November 1989, M. Zijlstra et al, "Germ-Line Transmission of a Disrupted $\beta$ 2-Microglobulin Gene Produced by Homologous Recombination in Embryonic Stem Cells", pages 435-438, especially page 436, cols. 1 and 2 to page 438, col. 1, parag. 1 and 2)	1-21
Y	Gene, Volume 85, issued 1989, K. Paludan et al, "Graduated Resistance of G418 Leads to Differential Selection of Cultured Mammalian Cells Expressing <u>neo</u> Gene" pages 421-426, especially abstract and page 422, col. 1, parag. 3 and col. 2, parag. 1, and page 424, col. 2, parag. 2.	1-16

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

- \* Special categories of cited documents:
- |   |  |
|---|--|
| *A* document defining the general state of the art which is not considered to be part of particular relevance   | *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  |
| *E* earlier document published on or after the international filing date  | *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone   |
| *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| *O* document referring to an oral disclosure, use, exhibition or other means  | *Z* document member of the same patent family  |
| *P* document published prior to the international filing date but later than the priority date claimed  |  |

Date of the actual completion of the international search

11 May 1993

Date of mailing of the international search report

17 MAY 1993

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Authorized officer

DEBORAH CROUCH, PH.D.

Facsimile No. NOT APPLICABLE

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)\*

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US93/00926

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Proceedings of the National Academy of Sciences, Volume 79, issued January 1982, B. Blomberg et al, "DNA Sequences of the Joining Regions of Mouse Lambda Light Chain Immunoglobulin Genes", pages 530-533, especially page 532, fig. 2.	21

**THIS PAGE BLANK (USPTO)**

---